Optimal solution of investment problems via linear parabolic equations generated by Kalman filter [*†]

Nikolai Dokuchaev

Department of Mathematics and Statistics, University of Limerick, Ireland

January 12, 2005

Abstract

We consider optimal investment problems for a diffusion market model with non-observable random drifts that evolve as an Itô's process. Admissible strategies do not use direct observations of the market parameters, but rather use historical stock prices. For a non-linear problem with a general performance criterion, the optimal portfolio strategy is expressed via the solution of a scalar minimization problem and a linear parabolic equation with coefficients generated by the Kalman filter.

Key words: *Optimal portfolio, non-observable parameters, Kalman filter*

Mathematical Subject Classification (1991): 49K45, 60G15, 93E20

Abbreviated title: *Optimal solution of investment problems*

1 Introduction

The paper investigates an optimal investment problem for a market which consists of a locally risk free asset, bond or bank account with interest rate $r(t)$, and a finite number, n, of risky stocks. We assume that the vector of stock prices $S(t)$ evolves according to an Itô stochastic differential equation $dS_i(t) = S_i(t)[a_i(t)\ dt + \sum_j \sigma_{ij}(t)\ dw_j(t)]$, $i = 1, \dots, n$, with a vector of appreciation rates $a(t)$ and a volatility matrix $\sigma(t)$. The problem goes back to Merton (1969), who found strategies which solve the optimization problem in which $\mathbf{E}U(X(T))$ is to be maximized, where $X(T)$ represents the wealth at the final time T and where $U(\cdot)$ is a utility function. If the market parameters are observed, then the optimal strategies (i.e. current vector of stock holdings) are functions of the current vector $(r(t), a(t), \sigma(t), S(t), X(t))$ (see, e.g., the

[*]Supported by NSERC of Canada under NCE Grant 30354 and Research Grant 88051.

[†]*SIAM J. of Control and Optimization* (2005) **44**, No. 4, pp. 1239-1258.

survey in Hakansson (1997) and Karatzas and Shreve (1998)). But in practice, $a(t)$ and $\sigma(t)$) have to be estimated from historical stock prices or some other observation process. There are many papers devoted to estimation of $(a(\cdot), \sigma(\cdot))$, mainly based on modifications of Kalman-Bucy filtering or the maximum likelihood principle (see e.g. Lo (1988), Chen and Scott (1993), Pearson and Sun (1994)). Unfortunately, the process $a(\cdot)$ is usually hard to estimate in real-time markets, because the drift term, $a(\cdot)$, is usually overshadowed by the diffusion term, $\sigma(\cdot)$. On the other hand, $\sigma(t)$ can, in principle, be found from stock prices. Thus, there remains the problem of optimal investment with unobservable $a(\cdot)$.

In fact, the problem is one of linear filtering. If $R_i(t)$ is the return on the ith stock, then $dR(t) = a(t)dt + \sigma(t)dw(t)$, so the estimation of $a(t)$ given $\{R(\tau), \tau < t\}$ (or $\{S(\tau), \tau < t\}$) is a linear filtering problem. If $a(\cdot)$ is conditionally Gaussian, then the Kalman filter provides the estimate which minimizes the error in the mean square sense.

A popular tool in optimal control and filtering theory is the separation theorem. This theorem has an analog in portfolio theory: it is the so-called "certainty equivalence principle": agents who know the solution of the optimal investment problem for the case of directly observable $a(t)$ can solve the problem with unobservable $a(t)$ by substituting $\mathbf{E}\{a(t)|S(\tau), \tau < t\}$ (see, e.g., Gennotte (1986)). Unfortunately, this principle does not hold in the general case of non-log utilities (see Kuwana (1995)). Note that this principle is unrelated to the much more recent notion of "certainty equivalent value" to be found in the work of Frittelli (2000).

Williams (1977), Detemple (1986), Dothan and Feldman (1986), Gennotte (1986), Brennan (1998) solved the investment problem using the Kalman-Bucy filter and dynamic programming. By this method, the optimal strategy can be calculated via solution of the Bellman parabolic equation; this equation is non-linear.

Karatzas (1997), Karatzas and Zhao (1998), Dokuchaev and Zhou (2000), Dokuchaev and Teo (2000) have obtained optimal portfolio strategies in general non-Gaussian setting, but only for case of time independent coefficients.

An approach based on Malliavin calculus gives a possibility to consider more general setting. Lakner (1995), (1998) assumes that $S(\cdot)$ and $w(\cdot)$ have equal dimension (as we do), and that $r(\cdot)$ and $\sigma(\cdot)$ are deterministic. This again guarantees that the filtration of $S(\cdot)$ is Brownian. Results from filtering theory give a representation of the optimal portfolio, which is explicit in terms of a conditional expectation of a Malliavin derivative when the $a_i(\cdot)$ are Ornstein-Uhlenbeck processes independent of $w(\cdot)$. Karatzas and Xue (1990) assume that there are more Brownian motions than stocks. They assume that $r(\cdot)$ and $\sigma(\cdot)$ are adapted to the observable $S(\cdot)$. After projecting onto an n-dimensional Brownian motion which generates the same filtration as $S(\cdot)$, they obtain a reduced, completely observable model; existence of an optimal portfolio follows, but the optimal strategy is, as usual, defined only implicitly.

We also consider the optimal investment problem with random and unobservable $a(\cdot)$. Following Lakner (1998) and Rishel (1999), we assume that $a(t)$ is a Gaussian process modelled by a system of linear Itô's equations. However, we consider a more general case when $(a(\cdot), r(\cdot))$ may depend on the realized returns (i.e., $b(\cdot) \neq 0$ in equation (2.4) below, and $r(\cdot)$ is correlated with $S(\cdot)$). We express the optimal strategy via solution of a Cauchy problem (4.3),(4.8) for a *linear* parabolic equation in $(n+1)$-dimensional vector space. Thus, we propose a simpler method than dynamic programming: the *nonlinear* parabolic Bellman equation is replaced for a linear parabolic equation. Note that the solution in Lakner (1998) expresses the optimal strategy via a conditional expectation of a random claim that depends on $w(\cdot)$; the solution presented below is also based on the martingale method but it is more constructive provided we can solve the Cauchy problem (4.3),(4.8). Using the technique of backward stochastic partial differential equations, we prove existence and uniqueness of the solution for this Cauchy problem. Furthermore, the most restrictive condition in Lakner (1998) was that the initial covariance of $a(0)$ is small enough (the condition (3.5) below). We replace it by another condition (4.9) that depends on U: it is less restrictive than (3.5) for some U's and more restrictive for others U's. For some problems, our condition (4.9) is automatically satisfied. In addition, we allow correlated $a(\cdot)$ and $w(\cdot)$.

2 The Model and Definitions

Consider a diffusion model of a market consisting of a locally risk free bank account or bond with price $B(t)$, $t \geq 0$, and n risky stocks with prices $S_i(t)$, $t \geq 0$, $i = 1, 2, \ldots, n$, where $n < +\infty$ is given. The prices of the stocks evolve according to the following equations:

$$dS_i(t) = S_i(t)\left(a_i(t)dt + \sum_{j=1}^{n} \sigma_{ij}(t)dw_j(t)\right), \quad t > 0, \tag{2.1}$$

where $w_i(t)$ are standard independent Wiener processes, $a_i(t)$ are appreciation rates, and $\sigma_{ij}(t)$ are volatility coefficients. The initial price $S_i(0) > 0$ is a given non-random constant. The price of the bond evolves according to the following equation

$$B(t) = B(0)\exp\left(\int_0^t r(t)dt\right), \tag{2.2}$$

where $B(0)$ is a given constant which we take to be 1 without loss of generality, and $r(t)$ is the random process of the risk-free interest rate.

We are given a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, where Ω is a set of elementary events, $\mathcal{F}$ is a complete σ-algebra of events, and $\mathbf{P}$ is a probability measure.

We introduce the vector processes ($^\top$ denoted transpose)

$$w(t) = (w_1(t), \ldots, w_n(t))^\top, \quad S(t) = (S_1(t), \ldots, S_n(t))^\top, \quad a(t) = (a_1(t), \ldots, a_n(t))^\top,$$

and the matrix process $\sigma(t) = \{\sigma_{ij}(t)\}_{i,j=1}^n$.

Let $\mathbf{1} \triangleq (1, \ldots, 1)^\top \in \mathbf{R}^n$, and $\widetilde{a}(t) \triangleq a(t) - r(t)\mathbf{1}$.

We define the return to time t by $dR_i(t) = dS_i(t)/S_i(t)$, $R_i(0) = 0$, and introduce the vector of returns $R(t) = (R_1(t), \ldots, R_n(t))^\top$ and of excess returns $\widetilde{R}_i(t) = R_i(t) - \int_0^t r(s)\,ds$.

Let $\{\mathcal{F}_t^{S,r}\}_{0 \le t \le T}$ be the filtration generated by the process $(r(t), S(t))$ completed with the null sets of $\mathcal{F}$.

Set $\widetilde{S}(t) \triangleq \exp\left(-\int_0^t r(s)ds\right) S(t)$.

We denote by $|x|$ the Euclidean norm of a vector $x \in \mathbf{R}^k$. For an Euclidean space E, we denote by $\mathrm{B}([0,T];E)$ the set of bounded measurable functions $f(t) : [0,T] \to E$. We denote by I_n the identity matrix in $\mathbf{R}^{n \times n}$. As usual, we say that $A < B$ for symmetric matrices if the matrix $B - A$ is definitely positive. We denote $\phi^- \triangleq \max(0, -\phi)$, and we denote by $\mathbb{I}_{\{\cdot\}}$ the indicator function.

The model for r, σ, and a

To describe the distribution of $\widetilde{a}(t)$, we shall use the model introduced in Lakner (1998, p.84), generalized for our case of random r, non-constant coefficients for the equation for $\widetilde{a}$, and correlated r, $\widetilde{a}$, and w. We assume that we are given measurable deterministic processes $\alpha(t)$, $\beta(t)$, $b(t)$ and $\delta(t)$ such that

$$d\widetilde{a}(t) = \alpha(t)[\delta(t) - \widetilde{a}(t)]dt + b(t)d\widetilde{R}(t) + \beta(t)dW(t), \tag{2.3}$$

where $\alpha(t) \in \mathbf{R}^{n \times n}$, $\beta(t) \in \mathbf{R}^{n \times n}$, $b(t) \in \mathbf{R}^{n \times n}$, $\delta(t) \in \mathbf{R}^n$, and where W is an n-dimensional Wiener process in $(\Omega, \mathcal{F}, P)$. We assume that $\alpha(t)$, $\beta(t)$, $b(t)$, and $\delta(t)$ are continuous in t and such that the matrix $\beta(t)$ is invertible and $|\beta(t)^{-1}| \le c$, where $c > 0$ is a constant. Further, we assume that $\widetilde{a}(0)$ follows an n-dimensional normal distribution with mean vector m_0 and covariance matrix γ_0. The vector m_0 and the matrix γ_0 are assumed to be known. We note that this setting covers the case when $\widetilde{a}$ is an n-dimensional Ornstein-Uhlenbeck process with mean-reverting drift.

Clearly, equation (2.3) can be rewritten as

$$d\widetilde{a}(t) = \Big(\alpha(t)\delta(t) + [b(t) - \alpha(t)]\widetilde{a}(t)\Big)dt + b(t)\sigma(t)dw(t) + \beta(t)dW(t). \tag{2.4}$$

In addition, it can be seen that $\widetilde{R}_i(t)$ evolve as

$$d\widetilde{R}_i(t) = \widetilde{a}_i(t)dt + \sum_{j=1}^{n} \sigma_{ij}(t)dw_j(t), \quad t > 0. \tag{2.5}$$

We assume that the process $\sigma(t)$ is continuous in t, non-random and such that $\sigma(t)\sigma(t)^\top \geq c_\sigma I_n$, where $c_\sigma > 0$ is a constant.

Further, we assume that $r(\cdot) = \phi_r(\widetilde{R}(\cdot), \Theta)$, where Θ is a random element in a metric space $\mathcal{X}_r$, and where $\phi_r : C([0,T]; \mathbf{R}^n) \times \mathcal{X}_r \to \mathrm{B}([0,T]; \mathbf{R})$ is a measurable function, and Θ does not depend on $(w(\cdot), W(\cdot), \widetilde{a}(0))$. In addition, we assume that the process $r(t)$ is adapted to the filtration generated by $(\widetilde{R}(t), \Theta)$. Note that closed system (2.4)-(2.5) for the pair $(\widetilde{a}(t), \widetilde{R}(t))$ does not include $r(\cdot)$, and $(\widetilde{a}(\cdot), \widetilde{R}(\cdot))$ does not depend on Θ. Therefore, the market model is well defined. The assumptions for measurability of r don't look very natural. However, they cover generic models when r is independent on $\widetilde{R}$ or non-random, and we can still consider some models with correlated r and $\widetilde{R}$.

Under these assumptions, the solution of (2.1) is well defined, but the market is incomplete.

Let $\widetilde{\phi}_m(t,s)$, $m = 0, 1$, be the solution of the matrix equation

$$\begin{cases} \frac{d\widetilde{\phi}_m}{dt}(t,s) = [m \cdot b(t) - \alpha(t)]\widetilde{\phi}_m(t,s), \\ \widetilde{\phi}_m(s,s) = I_n. \end{cases}$$

Let

$$\widetilde{K}_m(t) \triangleq \int_0^t \widetilde{\phi}_m(t,s) b(s)\sigma(s)\sigma(s)^\top b(s)^\top \widetilde{\phi}_m(t,s)^\top ds, \quad m = 0, 1. \tag{2.6}$$

We have that

$$\widetilde{a}(t) = \widetilde{\phi}_1(t,0)\widetilde{a}(0) + \int_0^t \widetilde{\phi}_1(t,s)[\alpha(s)\delta(s)ds + b(s)\sigma(s)dw(s) + \beta(s)dW(s)].$$

It follows that $\widetilde{K}_1(t)$ is the covariance matrix for $\widetilde{a}(t)$ calculated with $\beta(t) \equiv 0$ and $\widetilde{a}(0) = 0$. By the linearity of (2.4), it follows that $\widetilde{K}_1(t)$ is the conditional covariance for $\widetilde{a}(t)$ given $(W(\cdot)|_{[0,t]}, \widetilde{a}(0))$ or $(W(\cdot)|_{[0,T]}, \widetilde{a}(0))$.

Note that $\widetilde{K}_m(t)$ can be found as solutions of linear equations that one can easy derive from (2.4) and (4.1) (see, e.g., Arnold (1973), Chapter 8).

We assume that b is "small". More precisely, we assume that there exists $\varepsilon > 0$ such that

$$T\widetilde{K}_m(t) + \varepsilon I_n < \sigma(t)\sigma(t)^\top \quad \forall t \in [0,T], \quad m = 0,1. \tag{2.7}$$

The risk neutral probability measure

Set $Q(t) \triangleq (\sigma(t)\sigma(t)^\top)^{-1}$, and set

$$\mathcal{Z} \triangleq \exp\left(\int_0^T [\sigma(t)^{-1}\widetilde{a}(t)]^\top dw(t) + \frac{1}{2}\int_0^T \widetilde{a}(t)^\top Q(t)\widetilde{a}(t)dt\right). \tag{2.8}$$

Proposition 2.1

$$\mathbf{E}\Big\{\exp\frac{1}{2}\int_0^T \widetilde{a}(t)^\top Q(t)\widetilde{a}(t)dt \,\Big|\, W(\cdot), \widetilde{a}(0)\Big\} < +\infty \qquad a.s. \tag{2.9}$$

By this proposition, the Novikov's condition is satisfied conditionally, $\mathbf{E}\{\mathcal{Z}^{-1} \,|\, W(\cdot), \widetilde{a}(0)\} = 1$, then $\mathbf{E}\mathcal{Z}^{-1} = 1$.

Define the (equivalent) probability measure $\mathbf{P}_*$ by $d\mathbf{P}_*/d\mathbf{P} = \mathcal{Z}^{-1}$. Let $\mathbf{E}_*$ be the corresponding expectation.

The wealth and strategies

Let $X_0 > 0$ be the initial wealth at time $t = 0$, and let $X(t)$ be the wealth at time $t > 0$, $X(0) = X_0$. We assume that

$$X(t) = \pi_0(t) + \sum_{i=1}^n \pi_i(t), \tag{2.10}$$

where the pair $(\pi_0(t), \pi(t))$ describes the portfolio at time t. The process $\pi_0(t)$ is the investment in the bond, $\pi_i(t)$ is the investment in the ith stock, $\pi(t) = (\pi_1(t), \ldots, \pi_n(t))^\top$, $t \geq 0$.

Definition 2.1 *The process* $\widetilde{X}(t) \triangleq \exp\left(-\int_0^t r(s)ds\right) X(t)$ *is called the normalized (or discounted) wealth.*

Let $\mathbf{S}(t) \triangleq \operatorname{diag}(S_1(t), \ldots, S_n(t))$ and $\widetilde{\mathbf{S}}(t) \triangleq \operatorname{diag}(\widetilde{S}_1(t), \ldots, \widetilde{S}_n(t))$ be diagonal matrices with the corresponding diagonal elements. The portfolio is said to be self-financing, if

$$dX(t) = \pi(t)^\top \mathbf{S}(t)^{-1} dS(t) + \pi_0(t)r(t)dt = \pi(t)^\top dR(t) + \pi_0(t)r(t)dt. \tag{2.11}$$

It follows from (2.10) that for such portfolios

$$\begin{aligned} dX(t) &= r(t)X(t)\,dt + \pi(t)^\top\left(\widetilde{a}(t)\,dt + \sigma(t)\,dw(t)\right), \\ d\widetilde{X}(t) &= B(t)^{-1}\pi(t)^\top d\widetilde{R}(t), \end{aligned} \tag{2.12}$$

so π alone suffices to specify the portfolio; the process π_0 is uniquely defined by π via (2.10),(2.12); π it is called a self-financing strategy.

Definition 2.2 *Let $\bar{\Sigma}$ be the class of all $\mathcal{F}_t^{S,r}$-predictable processes $\pi(\cdot)$ such that*

- $\int_0^T \left(|\pi(t)^\top \widetilde{a}(t)|^2 + |\pi(t)^\top \sigma(t)|^2\right) dt < \infty$ *a.s.*
- *there exists a constant q_π such that* $\mathbf{P}\left(\widetilde{X}(t) - X_0 \geq q_\pi, \forall t \in [0,T]\right) = 1.$

A process $\pi(\cdot) \in \bar{\Sigma}$ is said to be an *admissible* strategy with corresponding wealth $X(\cdot)$.

For an admissible strategy $\pi(\cdot)$, $X(t,\pi(\cdot))$ denotes the corresponding total wealth, and $\widetilde{X}(t,\pi(\cdot))$ the corresponding normalized total wealth. It follows that $\widetilde{X}(t,\pi(\cdot))$ is a $\mathbf{P}_*$-supermartingale with $\mathbf{E}_*\widetilde{X}(t,\pi(\cdot)) \leq X_0$ and $\mathbf{E}_*|\widetilde{X}(t,\pi(\cdot))| \leq |X_0| + 2|q_\pi|$.

Note that by definition, admissible strategies from $\bar{\Sigma}$ use observations of $r(t)$ and $S(t)$ only. For these strategies, the processes $X(t)$ and $\widetilde{X}(t)$ are $\mathcal{F}_t^{S,r}$-adapted.

The following definition is standard.

Definition 2.3 *Let ξ be a given random variable. An admissible strategy $\pi(\cdot)$ is said to replicate the claim ξ if $X(T,\pi(\cdot)) = \xi$ a.s.*

3 Problem statement and preliminary results

Let $T > 0$, let $\widehat{D} \subset \mathbf{R}$ be convex and bounded below, and let $X_0 \in \widehat{D}$ be given. Let $U(\cdot) : \widehat{D} \to \mathbf{R} \cup \{-\infty\}$ be such that $U(X_0) > -\infty$.

We may state our general problem as follows: Find an admissible self-financing strategy $\pi(\cdot)$ which solves the following optimization problem:

$$\text{Maximize} \quad \mathbf{E}U(\widetilde{X}(T,\pi(\cdot))) \quad \text{over} \quad \pi(\cdot) \in \bar{\Sigma} \tag{3.1}$$

$$\text{subject to} \begin{cases} \widetilde{X}(0,\pi(\cdot)) = X_0, \\ \widetilde{X}(T,\pi(\cdot)) \in \widehat{D} \quad \text{a.s.} \end{cases} \tag{3.2}$$

The condition $\widetilde{X}(T,\pi(\cdot)) \in \widehat{D}$ may represent a requirement for a minimal normalized terminal wealth if $\widehat{D} = [k, +\infty)$, $k > 0$. This condition may represent also a requirement for the normalized terminal wealth in goal achieving problems if $\widehat{D} = [k_0, k_1]$, $k_0 < k_1$.

We assume that U, X_0 and $\widehat{D}$ satisfy the following two conditions.

Condition 3.1 *There exists a measurable set $\Lambda \subseteq [0,\infty)$, and a measurable function $F(\cdot,\cdot) : (0,\infty) \times \Lambda \to \widehat{D}$ such that for each $z > 0$, $\widehat{x} = F(z,\lambda)$ is a solution of the optimization problem*

$$\text{Maximize} \quad zU(x) - \lambda x \quad \text{over } x \in \widehat{D}. \tag{3.3}$$

Note that the usual concavity hypotheses imply this condition, but more general utility functions are also covered. For example, this condition is satisfied for the goal achieving problem when $U(x)$ is a step function (see e.g. Karatzas (1997), Dokuchaev and Zhou (2000)).

Let $\bar{\mathcal{Z}} \triangleq \mathbf{E}\{\mathcal{Z}|\mathcal{F}_T^{S,r}\}$. Since $(\widetilde{R}(\cdot), \widetilde{a}(\cdot))$ does not depend on Θ, we have that $\mathcal{Z}$ does not depend on Θ, and $\bar{\mathcal{Z}} = \mathbf{E}\{\mathcal{Z}|\widetilde{R}(\cdot)\}$. Let $F(\cdot)$ be as in Condition 3.1.

Condition 3.2 *There exists* $\widehat{\lambda} \in \Lambda$ *such that* $\mathbf{E}_*|F(\bar{\mathcal{Z}}, \widehat{\lambda})| < +\infty$ *and* $\mathbf{E}_* F(\bar{\mathcal{Z}}, \widehat{\lambda}) = X_0$.

We solve our problem in two steps using the martingale approach. First we show that $\mathbf{E}U(F(\bar{\mathcal{Z}}, \widehat{\lambda}))$ is an upper bound for the expected utility of normalized terminal wealth for $\pi(\cdot) \in \bar{\Sigma}$. Then we find a portfolio $\widehat{\pi}(\cdot)$ which replicates the claim $B(T)F(\bar{\mathcal{Z}}, \widehat{\lambda})$. This establishes the optimality of $\widehat{\pi}(\cdot)$.

The optimal claim

The following theorem is a reformulation of Theorem 2.5 from Lakner (1998) under slightly more general conditions that allow discontinuous functions F and U such as step functions.

Theorem 3.1 *(Dokuchaev and Haussmann (2000)). With* $\widehat{\lambda}$ *as in Condition 3.2, let* $\widehat{\xi} \triangleq F(\bar{\mathcal{Z}}, \widehat{\lambda})$. *Then*

(i) $\mathbf{E}\, U^-(\widehat{\xi}) < \infty$, $\widehat{\xi} \in \widehat{D}$ *a.s.;*

(ii) $\mathbf{E}U(\widehat{\xi}) \geq \mathbf{E}U(\widetilde{X}(T, \pi(\cdot)))$, $\forall \pi(\cdot) \in \bar{\Sigma}$*;*

(iii) The claim $B(T)\widehat{\xi}$ *is attainable in* $\bar{\Sigma}$, *and there exists a replicating strategy in* $\bar{\Sigma}$. *This strategy is optimal for problem (3.1)-(3.2).*

This theorem uses duality approach for constrained optimization that goes back to Lagrange, and $\widehat{\lambda}$ is the corresponding Lagrange multiplier.

Remark 3.1 Theorem 2.5 from Lakner (1998) was stated under some additional assumptions that can be formulated in our notations as
(i) $b(t) \equiv 0$, r is non-random, r, σ, α, β, δ are constant, and $\widehat{D} = (0, +\infty)$;
(ii) U is strictly concave and continuously differentiable on $(0, +\infty)$, and $\lim_{x \to +\infty} U'(x) = 0$;
(iii) there exists a function $J(\cdot) : \widehat{D} \to \mathbf{R}$ such that $J(\lambda/x) \equiv F(x, \lambda)$;
(iv) $\mathbf{E}_* J(\lambda/\bar{\mathcal{Z}}) < +\infty$ for any $\lambda > 0$.

Solution via conditional expectation

Let

$$\widehat{a}(t) \triangleq \mathbf{E}\{\widetilde{a}(t) \,|\, \mathcal{F}_t^{S,r}\}.$$

Set $\widetilde{\alpha}(t) \triangleq \alpha(t) - b(t)$ and $m_0 \triangleq \mathbf{E}\widetilde{a}(0)$.

Let $\gamma(t) \in \mathbf{R}^{n\times n}$ be the unique solution (in the class of symmetric nonnegative definite matrices) of the deterministic Riccati's equation

$$\begin{cases} \frac{d\gamma}{dt}(t) = -[b(t)\sigma(t)^\top + \gamma(t)]Q(t)[b(t)\sigma(t)^\top + \gamma(t)]^\top - \widetilde{\alpha}(t)\gamma(t) - \gamma(t)\widetilde{\alpha}(t)^\top + \beta(t)\beta(t)^\top, \\ \gamma(0) = \gamma_0. \end{cases} \tag{3.4}$$

Here $\gamma_0 \triangleq \mathbf{E}[\widetilde{a}(0) - m_0][\widetilde{a}(0) - m_0]^\top$. In fact, $\gamma(t) = \mathbf{E}\left\{[\widetilde{a}(t) - \widehat{a}(t)][\widetilde{a}(t) - \widehat{a}(t)]^\top|\mathcal{F}_t^{S,r}\right\}$.

Let $A(t) \triangleq -\widetilde{\alpha}(t) - \gamma(t)Q(t)$, and let $\phi(t)$ be the solution of the matrix equation

$$\begin{cases} \frac{d\phi}{dt}(t) = A(t)\phi(t), \\ \phi(0) = I_n, \end{cases}$$

where I_n is the unit matrix in $\mathbf{R}^{n\times n}$.

The following theorem is a reformulation of Theorem 4.3 from Lakner (1998). It gives the solution of the investment problem via conditional expectation of future values of some processes with known evolution.

Theorem 3.2 *(Lakner (1998)). Let conditions (i)-(iv) in Remark 3.1 holds, let $U(x)$ be twice differentiable on $(0, +\infty)$, and let*

$$\operatorname{tr}\gamma_0 + T\|\beta\|^2 < K_1, \quad K_1 = \frac{1}{360T\|\sigma^{-1}\|^2 K_0}, \quad K_0 = \max_{t\in[0,T]} \|e^{-\alpha t}\|^2, \tag{3.5}$$

where $\|\cdot\|$ denotes the Frobenius matrix norm, i.e., $\|\sigma^{-1}\|^2 = \operatorname{tr}[\sigma^{-1}\sigma^{-1\top}]$. Further, let

$$J(x) < K(1 + x^{-5}), \quad -J'(x) < K(1 - x^{-2}) \tag{3.6}$$

for some $K > 0$. Then the optimal strategy is

$$\pi(t)^\top = H(t)\bar{\mathcal{Z}}(t)\mathbf{E}\Big\{J'(\widehat{\lambda}\bar{\mathcal{Z}})\bar{\mathcal{Z}}^{-2}\Big[-\gamma(t)[\phi(t)^\top]^{-1}\int_t^T \phi(s)^\top[\sigma^\top]^{-1}d\widehat{w}(s) - \widehat{a}(t)\Big]\Big|\,\mathcal{F}_t^{S,r}\Big\},$$

where $H(t) \triangleq \widehat{\lambda}e^{r(t-T)}Q$ and $\widehat{w}(t) \triangleq w(t) - \int_0^t \sigma^{-1}\widehat{a}(s)ds$.

We propose below another solution such that the optimal strategy is presented via solution of a linear deterministic parabolic equation. We replace conditions (3.5) by condition (4.9) that can be less restrictive and is always satisfied if $\widehat{D}$ is bounded. In addition, we dropped condition (3.6) and the condition that (r, a) and w are independent: we allow $b(\cdot) \neq 0$ and $r = \phi_r(\widetilde{R}(\cdot), \Theta)$.

4 Main results: Solution via linear parabolic equation

Let $y(t)=(y_1(t),\ldots,y_{n+1}(t))=(\widehat{a}(t),y_{n+1}(t))$ be a process in $\mathbf{R}^{n+1}$, where

$$\begin{aligned}\widehat{a}(t) &= \mathbf{E}\{\widetilde{a}(t)|\mathcal{F}_t^{S,r}\},\\ y_{n+1}(t) &= -\frac{1}{2}\int_0^t \widehat{a}(s)^\top Q(s)\widehat{a}(s)ds+\int_0^t \widehat{a}(s)^\top Q(s)\,d\widetilde{R}(s).\end{aligned}$$

Let functions $f(\cdot):\mathbf{R}^{n+1}\times[0,T]\to\mathbf{R}^{n+1}$ and $g(\cdot):\mathbf{R}^{n+1}\times[0,T]\to\mathbf{R}^{(n+1)\times n}$ be such that

$$f(x,t)\triangleq\begin{pmatrix}[A(t)-b(t)\sigma(t)^\top Q(t)]\widehat{x}+\alpha(t)\delta(t)\\ -\frac{1}{2}\widehat{x}^\top Q(t)\widehat{x}\end{pmatrix},\qquad g(x,t)\triangleq\begin{pmatrix}[b(t)\sigma(t)^\top+\gamma(t)]Q(t)\\ \widehat{x}^\top Q(t)\end{pmatrix}.$$

Here $A(t)$ and $\gamma(t)$ are matrices defined above, $\gamma(t)$ is the solution of (3.4), and

$$x=(x_1,\ldots,x_{n+1})^\top=\begin{pmatrix}\widehat{x}\\ x_{n+1}\end{pmatrix},\qquad \widehat{x}=(x_1,\ldots,x_n)^\top.$$

By Theorem 10.3 from Liptser and Shiryaev (2000), p.396, the equation for $\widehat{a}(t)$ is

$$\begin{cases} d\widehat{a}(t)=[A(t)\widehat{a}(t)-b(t)\sigma(t)^\top Q(t)\widehat{a}(t)+\alpha(t)\delta(t)]dt+[b(t)\sigma(t)^\top+\gamma(t)]Q(t)d\widetilde{R}(t),\\ \widehat{a}(0)=m_0.\end{cases}\tag{4.1}$$

By (4.1)-(4.6), it follows that $y(\cdot)$ is the solution of the Itô's equation

$$\begin{cases} dy(t)=f(y(t),t)dt+g(y(t),t)\,d\widetilde{R}(t),\\ y(0)=y_0,\end{cases}\tag{4.2}$$

with

$$y_0=\begin{pmatrix}m_0\\ 0\end{pmatrix}\in\mathbf{R}^{n+1},\qquad m_0=\mathbf{E}\widetilde{a}(0).$$

The function $f(y,t)$ here does not satisfy Lipshitz condition with respect to $y\in\mathbf{R}^{n+1}$. However, the solution of this equation is uniquely defined. (It is shown in the proof of Lemma 4.1 below that the solution of (4.2) can be presented as a part of the unique solution of some Itô's equation with coefficients that are affine with respect to the state variable).

Lemma 4.1 *Let a function $\Phi(\cdot):\mathbf{R}^{n+1}\to\mathbf{R}$ be such that*

(i) $\mathbf{E}_*\Phi(y(T))=X_0$;

(ii) $\Phi(x)$ *is continuously twice differentiable;*

(iii) $\mathbf{E}_*\Phi(y(T))^2 < +\infty$.

Then there exists a unique classical solution $V : \mathbf{R}^{n+1} \times [0,T] \to \mathbf{R}$ *of the boundary value problem*

$$\frac{\partial V}{\partial t}(x,t) + \frac{\partial V}{\partial x}(x,t)f(x,t) + \frac{1}{2}\mathrm{tr}\left\{\frac{\partial^2 V}{\partial x^2}(x,t)\, g(x,t)\sigma(t)\sigma(t)^\top g(x,t)^\top\right\} = 0, \tag{4.3}$$

$$V(x,T) = \Phi(x). \tag{4.4}$$

Further, the processes $\widetilde{X}(t,\pi(\cdot)) \triangleq V(y(t),t)$ *and* $\pi(t)^\top \triangleq B(t)\frac{\partial V}{\partial x}(y(t),t)g(y(t),t)$*, are uniquely defined as elements of the spaces* $C([0,T], L_2(\Omega,\mathcal{F},P_*))$ *and* $L_2([0,T], L_2(\Omega,\mathcal{F},P_*))$ *respectively, and there exists a constant* $C > 0$ *such that*

$$\sup_{t\in[0,T]} \mathbf{E}_*|\widetilde{X}(t,\pi(\cdot))|^2 + \mathbf{E}_*\int_0^T B(t)^{-2}|\pi(t)|^2 dt \le C\mathbf{E}_*|\Phi(y(T))|^2 \tag{4.5}$$

for all these Φ*. Furthermore, the strategy* $\pi(t) = (\pi_1(t),\dots,\pi_n(t))$ *belongs to* $\bar{\Sigma}$ *and replicates the claim* $B(T)\Phi(y(T))$ *given the initial wealth* X_0 *with the normalized wealth* $\widetilde{X}(t) = V(y(t),t)$.

Note that estimate (4.5) reminds the Krylov-Ficera estimate (see Theorem 5.3.3 from Rozovskii (1980)) or its modification from Dokuchaev (1995)).

Further, we have that

$$d\bar{\mathcal{Z}}(t) = \widehat{a}(t)\bar{\mathcal{Z}}(t)d\widetilde{R}(t). \tag{4.6}$$

This formula (4.6) was derived in Theorem 3.1 from Lakner (1998) for the case when σ is constant and $b = 0$. The proof for a non-constant $\sigma(t)$ and $b \neq 0$ can be found in Dokuchaev and Haussmann (2000) and in Chapter 9 from Dokuchaev (2002). It follows that

$$y_{n+1}(t) = \ln \bar{\mathcal{Z}}(t). \tag{4.7}$$

Introduce the function $e(\cdot) : \mathbf{R}^{n+1} \to \mathbf{R}$ such that $e(y) = \exp[y_{n+1}]$ for $y = (y_1,\dots,y_{n+1})^\top$. Note that $\bar{\mathcal{Z}} = e(y(T))$.

Let $V = V(x,t,\lambda) : \mathbf{R}^{n+1} \times [0,T] \times \Lambda \to \mathbf{R}$ be the solution of partial differential equation (4.3) with the condition

$$V(x,T,\lambda) = F(e(x),\lambda). \tag{4.8}$$

The following result now is immediate.

Theorem 4.1 *Let* $\widehat{\lambda}$ *be such as in Condition 3.2. Assume that the function* $F(\cdot,\widehat{\lambda}) : \mathbf{R} \to \mathbf{R}$ *is such that is such that conditions (i)-(ii) of Lemma 4.1 are satisfied with* $\Phi(x) \triangleq F(e(x),\widehat{\lambda})$*, and*

$$\mathbf{E}_*F(\bar{\mathcal{Z}},\widehat{\lambda})^2 < +\infty. \tag{4.9}$$

Then there exists an unique classical solution V of problem (4.3)-(4.8) for $\lambda = \widehat{\lambda}$, and there exists an admissible self-financing strategy $\pi(\cdot) \in \bar{\Sigma}$ which replicates the claim $B(T)F(\bar{\mathcal{Z}}, \widehat{\lambda})$. This strategy is an optimal solution of problem (3.1)-(3.2), and

$$\widehat{\pi}(t)^\top = B(t)\frac{\partial V}{\partial x}(y(t), t, \widehat{\lambda})b(y(t), t), \quad \widetilde{X}(t, \pi(\cdot)) = V(y(t), t, \widehat{\lambda}), \tag{4.10}$$

Note that it is possible that condition (4.9) is not satisfied but the optimal claim $F(\bar{\mathcal{Z}}, \widehat{\lambda})$ is still replicable in the class of strategies $\bar{\Sigma}$. For example, let $U(x) \equiv \log x$, $X_0 = 1$, and $(0, +\infty) \subseteq \widehat{D}$, then $\Lambda = (0, \infty)$, $F(z, \lambda) = z/\lambda$, $\widehat{\lambda} = 1$, and the strategy is $\pi(t)^\top = B(t)\widehat{a}(t)^\top \bar{\mathcal{Z}}(t)Q(t)$ is replicating (and optimal) even in the case when (4.9) is not satisfied.

5 Special cases

Note that conditions (3.5) were imposed in Lakner (1998) with the only purpose to ensure that

$$\mathbf{E}_* \bar{\mathcal{Z}}^5 < +\infty, \qquad \mathbf{E}_* \bar{\mathcal{Z}}^{-4} < +\infty. \tag{5.1}$$

Our condition (4.9) for examples (i)-(iii) listed below is satisfied if $\mathbf{E}_* \bar{\mathcal{Z}}^\mu < +\infty$ for some $\mu \in \mathbf{R}$. For example (i), condition (4.9) is less restrictive than (5.1) if $l < 5/2$ and more restrictive if $l > 5/2$. For example (ii), condition (4.9) is less restrictive than (5.1) if $l < 2$ and more restrictive if $l > 2$. For example (iii), condition (4.9) is always less restrictive than (5.1). These examples are from Dokuchaev and Haussmann (2001):

(i) *Power utility.* Assume $\widehat{D} = [0, +\infty)$, $X_0 > 0$, $U(x) = d^{-1}x^d$, where either $d \in (0, 1)$ or $d < 0$. Then $\Lambda = (0, \infty)$, $F(z, \lambda) = (z/\lambda)^l$, and $\widehat{\lambda} = X_0^{-1/l}(\mathbf{E}_* \bar{\mathcal{Z}}^l)^{1/l}$, where $l = 1/(1-d)$.

(ii) Assume $\widehat{D} = [0, +\infty)$, $U(x) = -x^d + x$, where $d = 1 + 1/l$, and $l > 0$ is an integer, $X_0 > d^{-l}$. Then $\Lambda = [0, \infty)$, $F(z, \lambda) = (1 + \lambda/z)^l d^{-l}$, $\widehat{\lambda}$ is a root of a polynomial of degree l.

(iii) *Mean-variance utility.* Assume $\widehat{D} = \mathbf{R}$, $U(x) = -kx^2 + cx$, where $k \in \mathbf{R}$ and $c \geq 0$, $X_0 > 0$, then $F(z, \lambda) = (c - \lambda/z)/(2k)$.

We present below some sufficient conditions that ensure $\mathbf{E}_* \bar{\mathcal{Z}}^\mu < +\infty$ and, therefore, can be useful for verifying (6.2) .

Let $\widetilde{K}(t)$ be the covariance for $\widetilde{a}(t)$ under the probability measure $\mathbf{P}_*$, and let $\widehat{K}(t)$ be the covariance for $\widehat{a}(t)$ under $\mathbf{P}_*$.

Lemma 5.1 *If $\mu \in [0, 1]$, then $\mathbf{E}_* \bar{\mathcal{Z}}^\mu < +\infty$. Let $\mu < 0$ or $\mu > 1$. Then $\mathbf{E}_* \bar{\mathcal{Z}}^\mu < +\infty$ if there exist $\varepsilon > 0$ and $p > 1$ such that at least one of the following conditions holds:*

(i) $\kappa(p)\widehat{K}(t) < \sigma(t)\sigma(t)^\top - \varepsilon I_n$ for $t \in [0, T]$, where $\kappa(p) \triangleq qT(\mu^2 p - \mu) > 0$ with $q \triangleq p(p-1)^{-1}$.

(ii) $\kappa(p)\widetilde{K}(t) < \sigma(t)\sigma(t)^\top - \varepsilon I_n$ *for* $t \in [0,T]$.

It follows from Proposition 7.2 below that $\widetilde{K}(t)$ and $\widehat{K}(t)$ are the covariances of the processes defined by (2.4) and (4.1) respectively with $\widetilde{R}(\cdot)$ replaced by $\widetilde{R}_*(\cdot)$. Thus, these covariances can be found as solutions of linear deterministic equations that one can easy derive from (2.4) and (4.1) (see, e.g., Arnold (1973), Chapter 8).

6 Case of discontinuous F

To proceed further, we shall need a special weighted L_2-space with a weight defined via some parabolic equation. First, we introduce the operator

$$\mathcal{M}(t)p \stackrel{\Delta}{=} -\sum_{i=1}^{n+1} \frac{\partial}{\partial x_i} \left(p(x) f_i(x,t)\right) + \frac{1}{2}\sum_{i,j=1}^{n+1} \frac{\partial^2}{\partial x_i \partial x_j} \left(p(x)\widehat{g}_{ij}(x,t)\right),$$

where $\widehat{g} \stackrel{\Delta}{=} g\sigma\sigma^\top g^\top$.

Let $\rho_i \in L_2(\mathbf{R}^{n+1}) \cap C^2(\mathbf{R}^{n+1})$, $i=1,2$, be given such that $\rho_i(x) > 0$ for all $x \in \mathbf{R}^{n+1}$ and $\int_{\mathbf{R}^{n+1}} \rho_i(x)dx = 1$.

We consider the following parabolic equation

$$\begin{cases} \frac{\partial p}{\partial t}(x,t) = \mathcal{M}(t)p(x,t) + \rho_1(x), \quad t \in [0,T], \\ p(x,0) = \rho_0(x). \end{cases} \tag{6.1}$$

This boundary value problem has the unique classical solution $p(x,t)$ that is continuous in $\mathbf{R}^{n+1} \times [0,T]$. Let

$$\rho(x) \stackrel{\Delta}{=} \min_{t\in[0,T]} p(x,t).$$

We have that

$$p(\cdot,t) = \mathrm{G}(t,0)\rho_0 + \int_0^t \mathrm{G}(t,s)\rho_1 ds,$$

where $\mathrm{G}(t,s)$ is the semigroup operator generated by (6.1) (with $\rho_1 \equiv 0$) and such that $\mathrm{G}(s,s)\rho_i \equiv \rho_i$. We have that $(\mathrm{G}(t,s)\rho_i)(x) > 0$ for $t \in [s, s+\varepsilon)$ for some $\varepsilon = \varepsilon(x,s) > 0$. Hence $p(x,t) > 0$ for all x,t, and $\rho(x) > 0$ for all $x \in \mathbf{R}^{n+1}$. We shall use this ρ as a weight function.

We have that $\rho \in L_2(\mathbf{R}^{n+1}) \cap L_1(\mathbf{R}^{n+1})$, since $|\rho(x)| \le |\rho_0(x)|$.

We introduce the weighted space $L_{2,\rho}(\mathbf{R}^{n+1})$ with the norm

$$\|u\|_{L_{2,\rho}(\mathbf{R}^{n+1})} \stackrel{\Delta}{=} \left(\int_{\mathbf{R}^{n+1}} \rho(x)|u(x)|^2 dx\right)^{1/2}.$$

We introduce the space $\mathcal{Y}_k$ of functions $u = \{u_i(x,t)\}_{i=1}^k : \mathbf{R}^{n+1} \times [0,T] \to \mathbf{R}^k$ with the norm

$$\|u\|_{\mathcal{Y}_k} \triangleq \Big(\sum_{i=1}^k \int_0^T \|u_i(\cdot,t)\|^2_{L_{2,\rho}(\mathbf{R}^{n+1})}\, dt\Big)^{1/2}.$$

Further, we introduce the space $\mathcal{W}^1$ of functions $u = u(x,t) : \mathbf{R}^{n+1} \times [0,T] \to \mathbf{R}$ with the norm

$$\|u\|_{\mathcal{W}^1} \triangleq \|u\|_{\mathcal{Y}_1} + \left\|\frac{\partial u}{\partial x}\, g\right\|_{\mathcal{Y}_n}.$$

Finally, we introduce the space $\mathcal{W}^1_C$ consisting of all functions $u(\cdot) \in \mathcal{W}^1$ such that $u(\cdot) \in C([0,T]; L_{2,\rho}(\mathbf{R}^{n+1}))$ with the norm

$$\|u\|_{\mathcal{W}^1_C} \triangleq \sup_{t\in[0,T]} \|u(\cdot,t)\|_{L_{2,\rho}(\mathbf{R}^{n+1})} + \|u\|_{\mathcal{W}^1}.$$

The above space is a Banach space, since the weighted space $L_{2,\rho}(\mathbf{R}^{n+1})$ is a Hilbert space.

In fact, the spaces $\mathcal{Y}_k$, $\mathcal{W}^1$, and $\mathcal{W}^1_C$, are the completions in the corresponding norms of the set of smooth functions $u : \mathbf{R}^{n+1} \times [0,T] \to \mathbf{R}^k$ or $u : \mathbf{R}^{n+1} \times [0,T] \to \mathbf{R}$ respectively that have finite support.

Theorem 6.1 *Let p be the solution of (6.1), and let $\mathcal{W}^1_C$ be the corrsponding space defined via the weight $\rho(x) = \min_{t\in[0,T]} p(x,t)$. Let $\Phi(\cdot) : \mathbf{R}^{n+1} \to \mathbf{R}$ be a measurable function such that*

$$\int_{\mathbf{R}^{n+1}} p(x,T)\Phi(x)^2 dx < +\infty. \tag{6.2}$$

Then boundary value problem (4.3)-(4.4) admits a unique solution $V \in \mathcal{W}^1_C$. Moreover, there exists a constant $C > 0$ independent on $\Phi(\cdot)$ and such that

$$\|V\|^2_{\mathcal{W}^1_C} \le C \int_{\mathbf{R}^{n+1}} p(x,T)\Phi(x)^2 dx. \tag{6.3}$$

Note that condition (6.2) allows discontinuous Φ.

Remark 6.1 The definition of $\mathcal{W}^1_C$ ensures that problem (4.3)–(4.4) can be stated in $\mathcal{W}^1_C$. The functions V and $(\partial V/\partial x)g$ are measurable and $L_{2,\rho}$-integrable. The equality in (4.4) is the equality for elements of the space $L_{2,\rho}(\mathbf{R}^{n+1})$, it is meaningful since $V(\cdot,t)$ is continuous in t in $L_{2,\rho}(\mathbf{R}^{n+1})$. The equality in (4.3) is the equality for elements of the dual space $\mathcal{W}^{1*}$, since all components of $\frac{\partial^2 V}{\partial x^2}(x,t)\, g(x,t)\sigma(t)\sigma(t)^\top g(x,t)^\top$ belong to $\mathcal{W}^{1*}$.

It follows from the proof of Theorem 6.1 below that $\|p(\cdot,T)\|_{L_1(\mathbf{R}^{n+1})} = 2$. Hence (6.2) is satisfied for any bounded Φ. In addition, it can be shown that $\|p(\cdot,T)\|_{L_2(\mathbf{R}^{n+1})} \le$

$C\sum_{i=1,2}\|\rho_i(\cdot)\|_{L_2(\mathbf{R}^{n+1})}$, where $C>0$ is a constant that does not depend on ρ_i. Therefore, (6.2) is satisfied for any $\Phi\in L_4(\mathbf{R}^{n+1})$.

Theorem 6.1 gives the possibility to present the optimal investment strategy via solution of (4.3)–(4.4) for the case of discontinuous F. An example is the goal-achieving problem, when $\widehat{D}=[0,\infty)$, $X_0\in(0,\alpha)$, and $U(x)=0$ if $0\le x<\alpha$, $U(x)=1$ if $x\ge\alpha$. Then $\Lambda=(0,\infty)$, $F(z,\lambda)=\alpha$ if $0<\lambda\le z/\alpha$, $F(z,\lambda)=0$ if $\lambda>z/\alpha$, and (6.2) holds for $\Phi(x)=F(e(x),\lambda)$ $(\forall\lambda)$.

7 Appendix: Proofs

Proof of Proposition 2.1 . By Jensen' inequality, it follows that

$$\begin{aligned}\mathbf{E}\Big\{\exp\tfrac{1}{2}\textstyle\int_0^T\widetilde{a}(t)^\top Q(t)\widetilde{a}(t)dt\,\Big|\,W(\cdot),\widetilde{a}(0)\Big\}&=\mathbf{E}\Big\{\exp\tfrac{1}{T}\textstyle\int_0^T\tfrac{T}{2}\widetilde{a}(t)^\top Q(t)\widetilde{a}(t)dt\,\Big|\,W(\cdot),\widetilde{a}(0)\Big\}\\&\le\tfrac{1}{T}\textstyle\int_0^T\mathbf{E}\Big\{\exp\tfrac{T}{2}\widetilde{a}(t)^\top Q(t)\widetilde{a}(t)dt\,\Big|\,W(\cdot),\widetilde{a}(0)\Big\}.\end{aligned}$$

We have for definitely positive matrices that if $A>B>0$ then $B^{-1}>A^{-1}$. By condition (2.7) with $m=1$, it follows that

$$\begin{aligned}\widetilde{K}_0(t)^{-1}\;&>T[\sigma(t)\sigma(t)^\top-\varepsilon I_n]^{-1}=TQ(t)[I_n-\varepsilon Q(t)]^{-1}\\&=TQ(t)\left[I_n+\textstyle\sum_{k=1}^{+\infty}\{\varepsilon Q(t)\}^k\right]>TQ(t)+T\varepsilon Q(t)^2>TQ(t)+M,\end{aligned}\tag{7.1}$$

where $M=M(\varepsilon)>0$ is a definitely positive constant matrix. Clearly, we can take $\varepsilon>0$ small enough to ensure convergency of the series in (7.1).

To complete the proof, we shall use the following fact. Let ξ be a Gaussian n-dimensional vector, $K_\xi\triangleq\mathbf{E}(\xi-\mathbf{E}\xi)(\xi-\mathbf{E}\xi)^\top>0$. It is known that the probability density function for ξ is $C\exp[-\frac{1}{2}(x-\mathbf{E}\xi)^\top K_\xi^{-1}(x-\mathbf{E}\xi)]$, where $C>0$ is a constant. It follows that $\mathbf{E}\exp(\frac{1}{2}\xi^\top P\xi)<+\infty$ for any matrix $P\in\mathbf{R}^{n\times n}$ such that $0<P<K_\xi^{-1}$. Then the proof follows from (7.1). $\square$

We introduce the process

$$\widetilde{R}_*(t)\triangleq\int_0^t\sigma(s)\,dw(s).$$

Let n-dimensional vector random process $\widetilde{a}_*(t)$ be defined as the solution of

$$d\widetilde{a}_*(t)=\Big(\alpha(t)\delta(t)-\alpha(t)\widetilde{a}_*(t)\Big)dt+b(t)d\widetilde{R}_*(t)+\beta(t)dW(t),\quad\widetilde{a}_*(0)=\widetilde{a}(0).$$

Set

$$\mathcal{Z}_*\triangleq\exp\left(\int_0^T[\sigma(t)^{-1}\widetilde{a}_*(t)]^\top dw(t)-\frac{1}{2}\int_0^T\widetilde{a}_*(t)^\top Q(t)\widetilde{a}_*(t)dt\right).\tag{7.2}$$

Proposition 7.1 *There exists a measurable function* $\psi:C([0,T];\mathbf{R}^n)\times\mathrm{B}([0,T];\mathbf{R}^n)\to\mathbf{R}$ *such that* $\mathcal{Z}_*=\psi(\widetilde{R}_*(\cdot),\widetilde{a}_*(\cdot))$ *and* $\mathcal{Z}=\psi(\widetilde{R}(\cdot),\widetilde{a}(\cdot))$.

Proof. Clearly, ψ is defined by

$$\log \mathcal{Z}_* = \int_0^T \widetilde{a}_*(t)^\top Q(t) \left(d\widetilde{R}_*(t) - \frac{1}{2}\widetilde{a}_*(t)dt \right). \tag{7.3}$$

□

Let $r_*(\cdot) \triangleq \phi_r(\widetilde{R}_*(\cdot), \Theta)$ and $B_*(t) \triangleq B(0)\exp\left(\int_0^t r_*(s)ds\right)$ (ϕ_r is defined in Section 2). Let

$$\bar{\mathcal{Z}}_* \triangleq \mathbf{E}\{\mathcal{Z}_* | \widetilde{R}_*(\cdot), r_*(\cdot)\}. \tag{7.4}$$

Let $\mathcal{T} \triangleq C([0,T];\mathbf{R}^n) \times \mathbf{R}^n$. Clearly, there exists a measurable mapping $\mathcal{A} : [0,T] \times C([0,T];\mathbf{R}^n) \times \mathcal{T} \to C([0,T];\mathbf{R}^n)$ such that $\widetilde{a}_*(t) = \mathcal{A}(t, \widetilde{R}_*(\cdot), W(\cdot), \widetilde{a}(0))$ and $\widetilde{a}(t) = \mathcal{A}(t, \widetilde{R}(\cdot), W(\cdot), \widetilde{a}(0))$.

We have that $\bar{\mathcal{Z}}_* = \mathbf{E}\{\mathcal{Z}_* | \widetilde{R}_*(\cdot)\} = \bar{\psi}(\widetilde{R}_*(\cdot))$ and

$$\bar{\mathcal{Z}}_* = \mathbf{E}\{\psi([\widetilde{R}_*(\cdot), \widetilde{a}_*(\cdot)] | \widetilde{R}_*(\cdot)\} = \mathbf{E}\{\psi[\widetilde{R}_*(\cdot), \mathcal{A}(\cdot, \widetilde{R}_*(\cdot), W(\cdot), \widetilde{a}(0))] | \widetilde{R}_*(\cdot)\}.$$

By Proposition 7.1, it follows that

$$\bar{\mathcal{Z}} = \mathbf{E}\{\psi[\widetilde{R}(\cdot), \widetilde{a}(\cdot)] | \widetilde{R}(\cdot)\} = \mathbf{E}\{\psi[\widetilde{R}(\cdot), \mathcal{A}(\cdot, \widetilde{R}(\cdot), W(\cdot), \widetilde{a}(0))] | \widetilde{R}(\cdot)\}.$$

Hence there exists a measurable mapping $\bar{\psi}(\cdot) : C([0,T];\mathbf{R}^n) \to \mathbf{R}$ such that

$$\bar{\mathcal{Z}} = \bar{\psi}(\widetilde{R}(\cdot)), \quad \bar{\mathcal{Z}}_* = \bar{\psi}(\widetilde{R}_*(\cdot)). \tag{7.5}$$

Proposition 7.2 *Let a function $\phi : C([0,T];\mathbf{R}^n) \times \mathrm{B}([0,T];\mathbf{R}^n) \times \mathrm{B}([0,T];\mathbf{R}) \to \mathbf{R}$ be such that $\mathbf{E}\phi^-(\widetilde{R}(\cdot), \widetilde{a}(\cdot), r(\cdot)) < +\infty$. Further, let a function $\widehat{\phi} : C([0,T];\mathbf{R}^n) \times \mathrm{B}([0,T];\mathbf{R}) \to \mathbf{R}$ be such that $\mathbf{E}\widehat{\phi}^-(\widetilde{R}(\cdot), r(\cdot)) < +\infty$. Then*

$$\mathbf{E}\phi(\widetilde{R}(\cdot), \widetilde{a}(\cdot), r(\cdot)) = \mathbf{E}\mathcal{Z}_*\phi(\widetilde{R}_*(\cdot), \widetilde{a}_*(\cdot), r_*(\cdot)), \tag{7.6}$$

$$\mathbf{E}\widehat{\phi}(\widetilde{R}(\cdot), r(\cdot)) = \mathbf{E}\bar{\mathcal{Z}}_*\widehat{\phi}(\widetilde{R}_*(\cdot), r_*(\cdot)), \tag{7.7}$$

$$\mathbf{E}_*\widehat{\phi}(\widetilde{R}(\cdot), r(\cdot)) = \mathbf{E}\widehat{\phi}(\widetilde{R}_*(\cdot), r_*(\cdot)). \tag{7.8}$$

Proof. By assumption $(\Theta, W(\cdot), \widetilde{a}(0))$ is independent of $w(\cdot)$. To prove (7.6) it suffices to prove

$$\mathbf{E}\left\{\phi(\widetilde{R}(\cdot), \widetilde{a}(\cdot), r(\cdot)) \Big| \Theta, W(\cdot), \widetilde{a}(0)\right\} = \mathbf{E}\left\{\mathcal{Z}_*\phi(\widetilde{R}_*(\cdot), \widetilde{a}_*(\cdot), r_*(\cdot)) \Big| \Theta, W(\cdot), \widetilde{a}(0)\right\} \quad \text{a.s.} \tag{7.9}$$

Thus, for the next paragraph, without loss of generality, we shall suppose that $(\Theta, W(\cdot), \widetilde{a}(0))$ is deterministic, since for each value of $(\Theta, W(\cdot), \widetilde{a}(0))$ we can construct $\widetilde{R}, \widetilde{R}_*, \widetilde{a}, \widetilde{a}_*$.

By the linearity of (2.4), it follows that $\widetilde{K}_0(t)$ defined by (2.6) is the conditional covariance for $\widetilde{a}_*(t)$ given $(W(\cdot), \widetilde{a}(0))$. Similarly to the proof of Proposition 2.1, it can be shown that (2.7) with $m = 0$ ensures that $\mathbf{E}\{\mathcal{Z}_*|\Theta, W(\cdot), \widetilde{a}(0)\} = 1$ and $\mathbf{E}\mathcal{Z}_* = 1$. We define the probability measure $\bar{\mathbf{P}}$ by $d\bar{\mathbf{P}}/d\mathbf{P} = \mathcal{Z}_*$. (Each value of $(\Theta, W(\cdot), \widetilde{a}(0))$ generates its own $\bar{\mathbf{P}}$). By Girsanov's Theorem, the process

$$\bar{w}(t) \triangleq w(t) - \int_0^t \sigma(s)^{-1}\widetilde{a}_*(s)ds$$

is a Wiener process under $\bar{\mathbf{P}}$. From this we obtain

$$\begin{aligned} d\widetilde{R}(t) &= \mathcal{A}(t, \widetilde{R}(\cdot), W(\cdot), \widetilde{a}(0))dt + \sigma(t)dw(t), \\ d\widetilde{R}_*(t) &= \mathcal{A}(t, \widetilde{R}_*(\cdot), W(\cdot), \widetilde{a}(0))dt + \sigma(t)d\bar{w}(t). \end{aligned}$$

Then for each value of $(\Theta, W(\cdot), \widetilde{a}(0))$ the processes $(\widetilde{R}(\cdot), \widetilde{a}(\cdot), r(\cdot))$ and $(\widetilde{R}_*(\cdot), \widetilde{a}_*(\cdot), r_*(\cdot))$ have the same distribution on the probability spaces defined by $\mathbf{P}$ and $\bar{\mathbf{P}}$ respectively, and (7.9), hence (7.6) follows.

Further, (7.7) follows by taking conditional expectation in (7.6). Finally, using Proposition 7.1 and (7.6),

$$\begin{aligned} \mathbf{E}_*\widehat{\phi}(\widetilde{R}(\cdot), r(\cdot)) &= \mathbf{E}\mathcal{Z}^{-1}\widehat{\phi}(\widetilde{R}(\cdot), r(\cdot)) = \mathbf{E}\psi(\widetilde{R}(\cdot), \widetilde{a}(\cdot))^{-1}\widehat{\phi}(\widetilde{R}(\cdot), r(\cdot)) \\ &= \mathbf{E}\mathcal{Z}_*\psi(\widetilde{R}_*(\cdot), \widetilde{a}_*(\cdot))^{-1}\widehat{\phi}(\widetilde{R}_*(\cdot), r_*(\cdot)) = \mathbf{E}\widehat{\phi}(\widetilde{R}_*(\cdot), r_*(\cdot)). \qquad \square \end{aligned}$$

We turn now to Theorem 3.1. Define $\widehat{\xi}_* \triangleq F(\bar{\mathcal{Z}}_*, \widehat{\lambda})$. It follows from (7.5) that if we define $\widetilde{\phi}$ by $\widehat{\xi} = \widetilde{\phi}(\widetilde{R}(\cdot))$, then $\widehat{\xi}_* = \widetilde{\phi}(\widetilde{R}_*(\cdot))$.

Proof of Theorem 3.1. Let us show that $\mathbf{E}U^-(\widehat{\xi}) < \infty$ so that $\mathbf{E}U(\widehat{\xi})$ is well defined. For $k = 1, 2, \ldots$, we introduce the random events

$$\Omega_*^{(k)} \triangleq \{-k \le U(\widehat{\xi}_*) \le 0\}, \quad \Omega^{(k)} \triangleq \{-k \le U(\widehat{\xi}) \le 0\},$$

along with their indicator functions, $\mathbb{I}_*^{(k)}$ and $\mathbb{I}^{(k)}$, respectively. The number $\widehat{\xi}_*$ provides the unique maximum of the function $\bar{\mathcal{Z}}_*U(\xi_*) - \widehat{\lambda}\xi_*$ over $\widehat{D}$, and $X_0 \in \widehat{D}$. By Proposition 7.2, we have, for all $k = 1, 2, \ldots$,

$$\begin{aligned} \mathbf{E}\mathbb{I}^{(k)}U(\widehat{\xi}) - \mathbf{E}\mathbb{I}_*^{(k)}\widehat{\lambda}\widehat{\xi}_* &= \mathbf{E}\mathbb{I}_*^{(k)}\left(\bar{\mathcal{Z}}_*U(\widehat{\xi}_*) - \widehat{\lambda}\widehat{\xi}_*\right) \ge \mathbf{E}\mathbb{I}_*^{(k)}\left(\bar{\mathcal{Z}}_*U(X_0) - \widehat{\lambda}X_0\right) \\ &= \mathbf{E}\mathbb{I}^{(k)}U(X_0) - \widehat{\lambda}X_0\mathbf{P}(\Omega_*^{(k)}) \ge -|U(X_0)| - |\widehat{\lambda}X_0| > -\infty. \end{aligned}$$

Furthermore, we have that $\mathbf{E}|\widehat{\xi}_*| = \mathbf{E}_*|\widehat{\xi}| < +\infty$. Hence $\mathbf{E}U^-(\widehat{\xi}) < \infty$.

Now observe that for any $\pi \in \bar{\Sigma}$ we can apply (7.7) and (7.8) to $U(\widetilde{X}^\pi(T))$ (and use (7.5)) to obtain

$$\begin{aligned} \mathbf{E}U(\widetilde{X}^\pi(T)) = \mathbf{E}_*\{\bar{\mathcal{Z}}U(\widetilde{X}^\pi(T))\} &\le \mathbf{E}_*\{\bar{\mathcal{Z}}U(\widetilde{X}^\pi(T)) - \widehat{\lambda}\widetilde{X}^\pi(T)\} + \widehat{\lambda}X_0 \\ &\le \mathbf{E}_*\{\bar{\mathcal{Z}}U(\widehat{\xi}) - \widehat{\lambda}\widehat{\xi}\} + \widehat{\lambda}X_0 = \mathbf{E}_*\bar{\mathcal{Z}}U(\widehat{\xi}) = \mathbf{E}U(\widehat{\xi}). \end{aligned}$$

Thus (ii) is satisfied.

Let us show (iii). Since σ is non-random, hence w-adapted, then $\widehat{\xi}_* = \widehat{\phi}(w(\cdot))$, where $\widehat{\phi}(\cdot) : \mathrm{B}([0,T];\mathbf{R}^n) \to \mathbf{R}$ is a measurable functions. By the martingale representation theorem,

$$\widehat{\xi}_* = \mathbf{E}\widehat{\xi}_* + \int_0^T f(t, w(\cdot)|_{[0,t]})^\top dw(t),$$

where $f(t,\cdot) : \mathrm{B}([0,t];\mathbf{R}^n) \to \mathbf{R}^n$ is a measurable function such that $\int_0^T |f(t,w(\cdot)|_{[0,t]})|^2 dt < +\infty$ a.s. There exists a unique measurable function $f_0(t,\cdot) : \mathrm{B}([0,t];\mathbf{R}^n) \to \mathbf{R}^n$ such that $f(t,w(\cdot)|_{[0,t]}) \equiv f_0(t,\widetilde{R}_*(\cdot)|_{[0,t]})$. Thus,

$$\widehat{\xi}_* = \mathbf{E}\widehat{\xi}_* + \int_0^T f_0(t,\widetilde{R}_*(\cdot)|_{[0,t]})^\top dw(t) = \mathbf{E}\widehat{\xi}_* + \int_0^T f_0(t,\widetilde{R}_*(\cdot)|_{[0,t]})^\top \sigma(t)^{-1} d\widetilde{R}_*(t).$$

Proposition 7.2 implies that $\mathbf{E}\widehat{\xi}_* = \mathbf{E}_*\widehat{\xi} = X_0$, and

$$\widehat{\xi} = X_0 + \int_0^T f_0(t,\widetilde{R}(\cdot)|_{[0,t]})^\top \sigma(t)^{-1} d\widetilde{R}(t).$$

Hence the strategy $\widehat{\pi}(t)^\top = B(t) f_0(t,\widetilde{R}(\cdot)|_{[0,t]})^\top \sigma(t)^{-1}$ replicates $B(T)\widehat{\xi}$. It belongs to $\bar{\Sigma}$; in particular, since w and $\widetilde{R}$ generate the same sigma-algebra and $\widehat{D}$ is convex, then $\widetilde{X}(t,\pi(\cdot)) = \mathbf{E}\left\{\widehat{\xi}\,|\,\widetilde{R}(\cdot)|_{[0,t]}\right\} \in \widehat{D}$, hence bounded below. This completes the proof of Theorem 3.1. $\square$

Proof of Lemma 4.1. Let $\mathcal{V} \triangleq \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^{\frac{n(n+1)}{2}}$. Clearly, $\mathcal{V}$ is a $\widetilde{n}$-dimensional linear vector space, where $\widetilde{n} \triangleq n + 1 + n(n+1)/2$. Let $\widetilde{y}(t) = (\widetilde{y}_1(t), \widetilde{y}_2(t), \widetilde{y}_3(t))$ be a process in $\mathcal{V}$ such that

$$\widetilde{y}(t) = (y(t), \widetilde{y}_3(t)) = (\widetilde{y}_1(t), \widetilde{y}_2(t), \widetilde{y}_3(t)) = \left(\widehat{a}(t), \ln \bar{\mathcal{Z}}(t), \widehat{a}(t)\widehat{a}(t)^\top\right).$$

The last equality is satisfied by (4.7). It can be seen that the equation for $\widetilde{y}(t)$ is linear:

$$\begin{aligned}
d\widetilde{y}_1(t) &= [\widehat{A}(t)\widetilde{y}_1(t) + v(t)]dt + E(t)\, d\widetilde{R}(t),\\
d\widetilde{y}_2(t) &= -\tfrac{1}{2}\mathrm{Tr}\{Q(t)\widetilde{y}_3(t)\}\, dt + \widetilde{y}_1(t)^\top Q(t)\, d\widetilde{R}(t),\\
d\widetilde{y}_3(t) &= [\widehat{A}(t)\widetilde{y}_3(t) + \widetilde{y}_3(t)^\top \widehat{A}(t)^\top + v(t)\widetilde{y}_2(t)^\top + \widetilde{y}_2(t)v(t)^\top + \tfrac{1}{2}\{E(t)\sigma(t)\sigma(t)^\top E(t)^\top\}]dt\\
&\qquad + E(t)\, d\widetilde{R}(t)\widetilde{y}_2(t)^\top + y_2(t)\, d\widetilde{R}(t)^\top E(t)^\top.
\end{aligned}$$

Here $\widehat{A}(t)$, $v(t)$, $E(t)$ are known deterministic functions in $\mathbf{R}^{n\times n}$, $\mathbf{R}^n$ and $\mathbf{R}^{n\times n}$ respectively. In particular, $\widehat{A}(t) = A(t) - b(t)\sigma(t)^\top Q(t)$. Thus, the equation for $\widehat{y}(t)$ can be rewritten as

$$\begin{cases} d\widetilde{y}(t) = \widetilde{f}(\widetilde{y}(t),t)dt + \sum_{i=1}^n \widetilde{g}_i(\widetilde{y}(t),t)\, d\widetilde{R}_i(t),\\ \widetilde{y}(0) = \widetilde{y}_0, \end{cases} \tag{7.10}$$

with $\widetilde{y}_0 = \left(m_0, 0, m_0 m_0^\top\right)$, and with some functions $\widetilde{f}(\widetilde{x},t) : \mathcal{V} \times [0,T] \to \mathcal{V}$ and $\widetilde{g}_i(\widetilde{x},t) : \mathcal{V} \times [0,T] \to \mathcal{V}$, $i = 1,\ldots,n$, that are affine in $\widetilde{x} \in \mathcal{V}$ with continuous in t coefficients. In particular, $\partial \widetilde{f}(\widetilde{x},t)/\partial \widetilde{x}$ and $\partial \widetilde{g}_i(\widetilde{x},t)/\partial \widetilde{x}$ depend only on t, and they are uniformly bounded. Hence (7.10) has an unique solution. Therefore, equation (4.2) has the unique solution $y(t)$.

Let $\widetilde{V}(\widetilde{x},t) \triangleq \mathbf{E}_* \Phi(\bar{y}^{\tilde{x},s}(T))$, where the process $\bar{y}^{\tilde{x},s}(\cdot)$ takes values in $\mathbf{R}^{n+1}$ and is such that $\widetilde{y}^{\tilde{x},s}(\cdot) = (\bar{y}^{\tilde{x},s}(\cdot), \widetilde{y}_3^{\tilde{x},s}(\cdot))$ is the solution of (7.10) given the initial condition $\widetilde{y}(s) = \widetilde{x} \in \mathcal{V}$. Then $\widetilde{V}(\widetilde{x},t)$ is the classical solution of the boundary value problem for the corresponding backward Kolmogorov's equation

$$\begin{cases} \frac{\partial \widetilde{V}}{\partial t}(\widetilde{x},t) + \mathcal{L}(t)\widetilde{V}(\widetilde{x},t) = 0, \quad t \in [0,T], \\ V(\widetilde{x},T) = \Phi(\widetilde{x}_1, \widetilde{x}_2), \end{cases} \tag{7.11}$$

where $\widetilde{x} = (\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3) \in \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^{\frac{n(n+1)}{2}}$, and where $\mathcal{L}(t)$ is the second order differential operator on functions $v : \mathcal{V} \to \mathbf{R}$ generated by the Markov process $\widetilde{y}(t)$.

Let $y^{x,s}(\cdot) = (y_1^{x,s}(\cdot), \ldots, y_{n+1}^{x,s}(\cdot))$ be the solution of (4.2), and let $V(x,t) \triangleq \mathbf{E}_* \Phi(y^{x,t}(T))$. Clearly,

$$\widetilde{y}^{\tilde{x},s}(t) \equiv (y^{x,s}(t), \widehat{y}^{x,s}(t)\widehat{y}^{x,s}(t)^\top) = \left(y_1^{x,s}(t), \ldots, y_n^{x,s}(t), y_{n+1}^{x,s}(t),\, \widehat{y}^{x,s}(t)\widehat{y}^{x,s}(t)^\top\right),$$

if

$$\begin{aligned} &\widetilde{x} = (x, \widetilde{x}_3) = (\widehat{x}, x_{n+1}, \widehat{x}\widehat{x}^\top), \quad \widehat{x} = (x_1, \ldots, x_n), \quad x = (x_1, \ldots, x_n, x_{n+1}) = (\widehat{x}, x_{n+1}) \in \mathbf{R}^{n+1}, \\ &\widetilde{x}_3 = \widehat{x}\widehat{x}^\top \in \mathbf{R}^{n(n+1)/2}, \quad \widehat{y}^{x,s}(\cdot) = (y_1^{x,s}(\cdot), \ldots, y_n^{x,s}(\cdot)). \end{aligned}$$

In that case, $V(x,t) \equiv \widetilde{V}(x_1, \widehat{x}_2, \widehat{x}_2\widehat{x}_2^\top)$, where $x = (\widehat{x}, x_{n+1})$, $\widehat{x} \in \mathbf{R}^n$. Therefore, $V(x,t)$ is the classical solution of problem (4.3)-(4.4).

Let $y_*(\cdot)$ denotes the solution of (4.2) with $\widetilde{R}(\cdot)$ replaced by $\widetilde{R}_*(\cdot) = \int_0^\cdot \sigma(t)\,dw(t)$.

Set $\widetilde{X}_*(t) \triangleq V(y_*(t),t)$. From (4.3) and Itô's Lemma, it follows that

$$\widetilde{X}_*(T) = \widetilde{X}_*(t) + \int_t^T B_*(s)^{-1}\pi_*(s)^\top d\widetilde{R}_*(s),$$

where $\pi_*(t)^\top \triangleq B_*(t)\frac{\partial V}{\partial x}(y_*(t),t)g(y_*(t),t)$. It follows that $\widetilde{X}_*(0) = V(y_*(0),0) = \mathbf{E}V(y_*(T),T) = X_0$ and

$$d\widetilde{X}_*(t) = B_*(t)^{-1}\pi_*(t)^\top d\widetilde{R}_*(t), \quad \widetilde{X}_*(T) = \Phi(y_*(T)). \tag{7.12}$$

Then $\widetilde{X}_*(t) = \widetilde{\psi}(t, \widetilde{R}_*)$ for some measurable $\widetilde{\psi}$, and the result follows if we observe that $\widetilde{X}(t) = \widetilde{\psi}(t, \widetilde{R})$ replicates the claim as desired for $\pi(t)^\top \triangleq B(t)\frac{\partial V}{\partial x}(y(t),t)g(y(t),t)$.

To continue, we require some a priori estimates. Let $\zeta_*(t) \triangleq B_*(t)^{-1}\sigma(t)^\top\pi_*(t)$.

We consider the conditional probability space given $(\Theta, W(\cdot), \widetilde{a}(0))$. With respect to the conditional probability space, it follows from (7.12) that

$$\begin{cases} d\widetilde{X}_*(t) = \zeta_*(t)^\top dw(t), \\ \widetilde{X}_*(T) = \Phi(y_*(T)). \end{cases} \tag{7.13}$$

By Proposition 2.2 El Karoui *et al* (1997), the (unique) solution $(\zeta_*(t), \widetilde{X}_*(t))$ of linear stochastic backward equation (7.13) is a process in $L_2([0,T], L_2(\Omega, \mathcal{F}, P)) \times C([0,T], L_2(\Omega, \mathcal{F}, P))$, and there exists a constant c_0, independent of $(\Phi(\cdot), \Theta, W(\cdot), \widetilde{a}(0))$ and such that

$$\begin{aligned} \sup_{t\in[0,T]} \mathbf{E}\left\{|\widetilde{X}_*(t)|^2 |\, \Theta, W(\cdot),\, \widetilde{a}(0)\right\} \quad &+ \mathbf{E}\left\{\int_0^T |\zeta_*(t)|^2 dt \,\middle|\, \Theta, W(\cdot),\, \widetilde{a}(0)\right\} \\ &\le c_0 \mathbf{E}\left\{\Phi(y_*(T))^2 |\, \Theta, W(\cdot),\, \widetilde{a}(0)\right\} \quad \text{a.s.} \end{aligned}$$

Hence

$$\sup_{t\in[0,T]} \mathbf{E}|\widetilde{X}_*(t)|^2 + \mathbf{E}\int_0^T B_*(t)^{-2} |\pi_*(t)|^2 dt \le c_1 \mathbf{E}\Phi(y_*(T))^2, \tag{7.14}$$

where $c_1 > 0$ is a constant that does not depend on $\Phi(\cdot)$. Then (4.5) follows. This completes the proof. $\square$

Proof of Theorem 4.1. Clearly, the equation for $y(t)$ is

$$\begin{cases} d\widehat{a}(t) = [A(t)\widehat{y}(t) - b(t)\sigma(t)^\top Q(t) + \alpha(t)\delta(t)]dt + \gamma(t)Q(t)\, d\widetilde{R}(t), \\ dy_{n+1}(t) = \frac{1}{2}\widehat{a}(t)^\top Q(t)\widehat{a}(t)dt - \widehat{a}(t)^\top Q(t)\, d\widetilde{R}(t). \end{cases}$$

As in the proof above, it can be shown that $\widetilde{X}(t) = V(y(t), t, \widehat{\lambda})$ is the solution of some equation (7.12), i.e. it is the normalized wealth. Then the proof follows. $\square$

Let $\mathcal{N}_2$ be the set of all Gaussian processes $\bar{a}(t) : [0,T] \times \Omega \to \mathbf{R}^n$ which are progressively measurable with respect to the filtration generated by $[a(0), w(t), W(t)]$ and such that $\mathbf{E}\int_0^T |\bar{a}(t)|^2 dt < +\infty$. For $\bar{a}(\cdot) \in \mathcal{N}_2$, let

$$Z(t, \bar{a}(\cdot)) \triangleq \exp\left[\int_0^t \bar{a}(s)^\top Q(s) d\widetilde{R}(s) - \frac{1}{2}\int_0^t \bar{a}(s)^\top Q(s)\bar{a}(s) ds\right].$$

Proposition 7.3 *Let $\bar{a}(\cdot) \in \mathcal{N}_2$, let $p \in (1, +\infty)$, and let $\mu \in \mathbf{R}$, $\mu < 0$ or $\mu > 1$. Let $\bar{K}(t)$ be the covariance matrix of $\bar{a}(t)$ under $\mathbf{P}_*$, and let $\kappa(p) \triangleq qT(\mu^2 p - \mu)$, where $q \triangleq p(p-1)^{-1}$. Let $\kappa(p)\bar{K}(t) < \sigma(t)\sigma(t)^\top - \varepsilon I_n$, where $\varepsilon > 0$ is a constant. Then $\mathbf{E}_* Z(t, \bar{a}(\cdot))^\mu < +\infty$.*

Proof of Proposition 7.3. If $\mu \in [0,1]$, then $\mathbf{E}_* Z(t, \bar{a}(\cdot))^\mu < +\infty$ (see Lakner (1998), p.93). Therefore, we can assume without loss of generality that $\mu < 0$ or $\mu > 1$. Clearly,

$$Z(t, \bar{a}(\cdot))^\mu = \exp\left[\mu \int_0^t \bar{a}(s)^\top Q(s) d\widetilde{R}(s) - \frac{\mu}{2}\int_0^t \bar{a}(s)^\top Q(s)\bar{a}(s) ds\right] = \zeta(t)\zeta_0(t),$$

where

$$\zeta(t) \triangleq \exp\left[\mu\int_0^t \bar a(s)^\top Q(s)d\widetilde R(s) - \frac{\mu^2 p}{2}\int_0^t \bar a(s)^\top Q(s)\bar a(s)ds\right],$$

and

$$\zeta_0(t) \triangleq \exp\left[\frac{\mu^2 p-\mu}{2}\int_0^t \bar a(s)^\top Q(s)\bar a(s)ds\right].$$

By Hölder inequality, $\mathbf{E}_*\mathcal{Z}^\mu \le [\mathbf{E}_*\zeta(T)^p]^{1/p}\,[\mathbf{E}_*\zeta_0(T)^q]^{1/q}$.

Similarly to the proof of Lemma A.1 from Lakner (1998), we have that $\mathbf{E}_*\zeta(T)^p < +\infty$ because $\zeta(t)^p$ is a positive local martingale with respect to $\mathbf{P}_*$, thus by Fatou's lemma it is a supermartingale.

By Jensen's inequality,

$$\begin{aligned}&\mathbf{E}_*\zeta_0(T)^q = \mathbf{E}_*\exp\left[q\tfrac{\mu^2p-\mu}{2}\textstyle\int_0^T \bar a(s)^\top Q(s)\bar a(s)ds\right]\\ &= \mathbf{E}_*\exp\left[\tfrac{1}{2T}\kappa(p)\textstyle\int_0^T \bar a(s)^\top Q(s)\bar a(s)ds\right] \le \tfrac{1}{T}\textstyle\int_0^T \mathbf{E}_*\exp\left[\tfrac12\kappa(p)\bar a(s)^\top Q(s)\bar a(s)\right]ds.\end{aligned}\tag{7.15}$$

Remind that $Q \triangleq (\sigma\sigma^\top)^{-1}$, and $\kappa(p) > 0$. Similarly to (7.1), we obtain

$$\begin{aligned}\bar K(t)^{-1} \quad &> \kappa(p)[\sigma(t)\sigma(t)^\top - \varepsilon I_n]^{-1} = \kappa(p)Q(t)[I_n - \varepsilon Q(t)]^{-1}\\ &= \kappa(p)Q(t)\left[I_n + \textstyle\sum_{k=1}^{+\infty}\{\varepsilon Q(t)\}^k\right] > \kappa(p)Q(t) + \kappa(p)\varepsilon Q(t)^2 > \kappa(p)Q(t) + M_1,\end{aligned}\tag{7.16}$$

where $M_1 = M_1(\varepsilon) > 0$ is a definitely positive constant matrix. (We can take $\varepsilon > 0$ small enough to ensure convergency.) Similarly to the proof of Proposition 2.1, it follows from (7.15), (7.16) that $\mathbf{E}_*\zeta_0(T)^q < +\infty$ and $\mathbf{E}_*Z(t,\bar a(\cdot))^\mu < +\infty$. $\square$

Proof of Lemma 5.1. If $\mu \in [0,1]$, then $\mathbf{E}_*\bar{\mathcal{Z}}^\mu < +\infty$ (see Lakner (1998), p.93). Therefore, we can assume without loss of generality that $\mu < 0$ or $\mu > 1$. Note that $\widehat a(\cdot) \in \mathcal{N}_2$. By Proposition 7.3, if (i) is satisfied then $\mathbf{E}_*\bar{\mathcal{Z}}^\mu < +\infty$.

Further, let (ii) be satisfied. Clearly, $\widetilde a(\cdot) \in \mathcal{N}_2$. By Proposition 7.3 again, $\mathbf{E}_*Z(T,\widetilde a(\cdot))^\mu < +\infty$. By (7.4), $\bar{\mathcal{Z}}_* = \mathbf{E}\{Z(T,\widetilde a_*(\cdot))|\widetilde R_*(\cdot), r_*(\cdot)\}$. Hence by Jensen's inequality $\mathbf{E}_*\bar{\mathcal{Z}}^\mu \le \mathbf{E}_*Z(T,\widetilde a(\cdot))^\mu < +\infty$. $\square$

Proof of Theorem 6.1. Let τ be a random variable that takes values in $[0,T]$ and such that $\mathbf{P}(\tau = 0) = 1/2$ and $\mathbf{P}(\tau \in (t_1,t_2]) = (t_2-t_1)/(2T)$ for $0 < t_1 < t_2 \le T$. Let $\eta_i \in L_2(\Omega,\mathcal{F},\mathbf{P},\mathbf{R}^{n+2})$ be random vectors such that they have the probability density functions $\rho_i(x)$, $i = 0,1$. We assume that $\tau,\eta_0,\eta_1,w,\Theta,W(\cdot),\widetilde a(0)$ are mutually independent.

Let $\eta \triangleq \eta_0\mathbb{I}_{\{\tau=0\}} + \eta_1\mathbb{I}_{\{\tau>0\}}$, and let $\eta_*(\cdot)$ be the solution of the Itô's equation

$$\begin{cases} d\eta_*(t) = f(\eta_*(t),t)dt + g(\eta_*(t),t)d\widetilde R_*(t), \quad t > \tau,\\ \eta_*(\tau) = \eta.\end{cases}\tag{7.17}$$

Equation (6.1) is the forward Kolmogorov's equation for the case when time of birth is distributed as τ, and the vector $\eta_*(t)$ has the conditional probability density function $p(x,t)/2$ in the sense that $\mathbf{P}(\eta_*(t)\in\Gamma,\ t\ge\tau)=1/2\int_\Gamma p(x,t)dx$ for any domain $\Gamma\subset\mathbf{R}^{n+1}$, where p is the solution of (6.1).

Note that we need random τ with the selected probability density on $(0,T]$ to generate the free term in parabolic equation (6.1).

Assume that $\Phi(\cdot)\in C^2(\mathbf{R}^{n+1})$ and it has finite support. Let $V(x,t)\triangleq\mathbf{E}_*\Phi(y^{x,t}(T))$, where $y^{x,s}(\cdot)$ is the solution of (4.2). Then $V(x,t)$ is the classical solution of problem (4.3)-(4.4). Set $\widetilde{Y}_*(t)\triangleq V(\eta_*(t),t)$. From (4.3) and Itô's Lemma, it follows that

$$\widetilde{Y}_*(T)=\widetilde{Y}_*(t)+\int_t^T B_*(s)^{-1}\varrho_*(s)^\top d\widetilde{R}_*(s),\quad \tau\le t\le T,$$

where $\varrho_*(t)^\top\triangleq B_*(t)\frac{\partial V}{\partial x}(\eta_*(t),t)g(\eta_*(t),t)$. Hence

$$d\widetilde{Y}_*(t)=B_*(t)^{-1}\varrho_*(t)^\top d\widetilde{R}_*(t),\quad \widetilde{Y}_*(T)=\Phi(\eta_*(T)). \tag{7.18}$$

To continue, we require some estimates. Let $\widehat{\zeta}_*(t)\triangleq B_*(t)^{-1}\sigma(t)^\top\varrho_*(t)$.

Consider the conditional probability space given $(\tau,\eta,\Theta,W(\cdot),\widetilde{a}(0))$. With respect to the conditional probability space, it follows from (7.18) that

$$\begin{cases} d\widetilde{Y}_*(t)=\widehat{\zeta}_*(t)^\top dw(t),\\ \widetilde{Y}_*(T)=\Phi(\eta_*(T)).\end{cases} \tag{7.19}$$

By Proposition 2.2 El Karoui *et al* (1997)) again, the (unique) solution $(\widehat{\zeta}_*(t),\widetilde{Y}_*(t))$ of stochastic backward equation (7.19) is a process in $L_2([\tau,T],L^2(\Omega,\mathcal{F},P))\times C([\tau,T],L^2(\Omega,\mathcal{F},P))$ given $(\tau,\eta,\Theta,W(\cdot),\widetilde{a}(0))$, and there exists a constant C_0 that is independent of $(\Phi(\cdot),\tau,\eta,\Theta,W(\cdot),\widetilde{a}(0))$, and such that

$$\begin{aligned}&\sup_{t\in[0,T]}\mathbf{E}\mathbb{I}_{\{t\ge\tau\}}\left\{|\widetilde{Y}_*(t)|^2|\,\tau,\eta,\Theta,W(\cdot),\widetilde{a}(0)\right\}+\mathbf{E}\left\{\int_0^T\mathbb{I}_{\{t\ge\tau\}}|\widehat{\zeta}_*(t)|^2dt\,\middle|\,\tau,\eta,\Theta,W(\cdot),\widetilde{a}(0)\right\}\\ &=\sup_{t\in[\tau,T]}\mathbf{E}\left\{|\widetilde{Y}_*(t)|^2\,|\,\tau,\eta,\Theta,W(\cdot),\widetilde{a}(0)\right\}+\mathbf{E}\left\{\int_\tau^T|\widehat{\zeta}_*(t)|^2dt\,\middle|\,\tau,\eta,\Theta,W(\cdot),\widetilde{a}(0)\right\}\\ &\qquad\le C_0\mathbf{E}\left\{\Phi(\eta_*(T))^2\,|\,\tau,\eta,\Theta,W(\cdot),\widetilde{a}(0)\right\}\quad\text{a.s.}\end{aligned}$$

Hence there exists a constant c_0, independent of $\Phi(\cdot)$ and such that

$$\sup_{t\in[0,T]}\mathbf{E}\mathbb{I}_{\{t\ge\tau\}}|\widetilde{Y}_*(t)|^2+\mathbf{E}\int_0^T\mathbb{I}_{\{t\ge\tau\}}B_*(t)^{-2}|\varrho_*(t)|^2dt\le c_0\mathbf{E}\Phi(\eta_*(T))^2. \tag{7.20}$$

Let $\Phi(\cdot)$ be a general measurable function satisfying the conditions specified in the theorem. Then, there exists a sequence $\{\Phi^{(i)}(\cdot)\}$, where $\Phi^{(i)}(\cdot) \in C^2(\mathbf{R}^{n+1})$ are such that they all have finite support and

$$\mathbf{E}|\Phi^{(i)}(\eta_*(T)) - \Phi(\eta_*(T))|^2 = \textstyle\int_{\mathbf{R}^{n+2}} p(x,T)|\Phi^{(i)}(x) - \Phi(x)|^2 dx \to 0 \quad \text{as} \quad i \to \infty. \tag{7.21}$$

Let $\widetilde{Y}_*^{(i)}(\cdot)$, $\varrho_*^{(i)}(\cdot)$, and $V^{(i)}(\cdot)$ be the corresponding processes and functions. Let

$$\Psi_{i,j} \triangleq \sup_{t\in[0,T]} \mathbf{E}\mathbb{I}_{\{t\geq\tau\}}|\widetilde{Y}_*^{(i)}(t) - \widetilde{Y}_*^{(j)}(t)|^2 + \mathbf{E}\int_0^T \mathbb{I}_{\{t\geq\tau\}} B_*(t)^{-2}|\varrho_*^{(i)}(t) - \varrho_*^{(j)}(t)|^2 dt.$$

By (7.20)-(7.21) and the linearity of (7.19), it follows that

$$\Psi_{i,j} \leq c_0\mathbf{E}|\Phi^{(i)}(\eta_*(T)) - \Phi^{(j)}(\eta_*(T))|^2 \to 0 \quad \text{as} \quad i,j \to \infty.$$

We have that $V^{(j)} \in \mathcal{W}_C^1$, since they are bounded together with their partial derivatives with respect to $x_1, \ldots, x_{n+1}$. Remind that $0 < \rho(x) \leq p(x,t)$ for all x, t. Furthermore, we have that

$$\begin{aligned} \Psi_{i,j} &= \sup_{t\in[0,T]} \int_{\mathbf{R}^{n+1}} p(x,t)|V^{(i)}(x,t) - V^{(j)}(x,t)|^2 dx \\ &+ \int_s^T dt \int_{\mathbf{R}^{n+1}} p(x,t)\left|\left[\frac{\partial V^{(i)}}{\partial x}(x,t) - \frac{\partial V^{(j)}}{\partial x}(x,t)\right] g(x,t)\right|^2 dx. \end{aligned}$$

Hence $\|V^{(i)} - V^{(j)}\|^2_{\mathcal{W}_C^1} \leq \Psi_{i,j} \to 0 \quad \text{as} \quad i,j \to \infty$. Therefore, $V^{(i)}$ is a Cauchy sequence in $\mathcal{W}_C^1$, and it has the limit V in $\mathcal{W}_C^1$. This V is the desired solution, and (6.3) is satisfied. This completes the proof. □

Note that it follows from the proof above that the sequences $\{\widetilde{Y}_*^{(i)}(\cdot)\}_{i=1}^\infty$ and $\{\varrho_*^{(i)}(\cdot)\}_{i=1}^\infty$ are Cauchy sequences in the spaces $C([\tau,T]; L^2(\Omega,\mathcal{F},\mathbf{P}\{\cdot\,|\,\tau\}))$ and $L_2([\tau,T]; L^2(\Omega,\mathcal{F},\mathbf{P}\{\cdot\,|\,\tau\}))$ respectively. Hence the corresponding limits $\widetilde{Y}_*(\cdot)$, $\varrho_*(\cdot)$ exist and belong to these spaces given τ.

This paper presents development of some results and ideas that grew up from our collaboration with Ulrich Haussmann during the author's stay at Pacific Institute for the Mathematical Sciences, Vancouver (see, e.g., Dokuchaev and Haussmann (2000)). The author wish to thank Prof. U. Haussmann for the support and useful discussion. The author also wishes to thank the anonymous referees for their insightful comments which greatly strengthened the paper.

References

L. Arnold, Stochastic differential equations. Theory and applications. Wiley-Inter-Science. New York, 1973.

M.J. Brennan, The role of learning in dynamic portfolio decisions, *European Finance Rev.* **1** (1998), 295-306.

R.-R. Chen, L. Scott, Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates, *J. Fixed Income* **4** (1993), 14-31.

J.B. Detemple, Asset pricing in an economy with incomplete information, *J. Finance* (1986) **41**, 369-382.

N.G. Dokuchaev, Probability distributions of Ito's processes: estimations for density functions and for conditional expectations of integral functionals. *Theory of Probability and its Applications* **39** (1995), N 4. 662-670.

N.G. Dokuchaev, U. Haussmann, Optimal portfolio selection and compression in an incomplete market *Quant. Finance* **1** (2001), 336-345.

N.G. Dokuchaev, U. Haussmann, Adaptive portfolio selection based on Historical Prices. Working paper. Presented in Quantitative Risk Management in Finance, Carnegie Mellon University, Pittsburgh. July 31 - August 5, 2000.

N.G. Dokuchaev, K.L. Teo, Optimal hedging strategy for a portfolio investment problem with additional constraints, *Dynamics of Continuous, Discrete and Impulsive Systems*, **7** (2000), 385-404.

N.G. Dokuchaev, X.Y. Zhou, Optimal investment strategies with bounded risks, general utilities, and goal achieving, *J. Mathematical Economics*, **35** (2000), 289-309.

N.G. Dokuchaev, *Dynamic portfolio strategies: quantitative methods and empirical rules for incomplete information.* Kluwer Academic Publishers, Boston, January 2002.

U. Dothan, D. Feldman, Equilibrium interest rates and multiperiod bonds in a partially observable economy, *J. Finance* **41** (1986), 369-382.

N. El Karoui, S. Peng, and M.C. Quenez, Backward stochastic differential equations in finance. *Mathematical Finance* **7** (1997), 1–71.

M. Frittelli, Introduction to a theory of value coherent with the no-arbitrage principle, *Finance and Stochastics*, **3** (2000), 275-298.

G. Gennotte, Optimal portfolio choice under incomplete information, *J. Finance* **41** (1986), 733-749.

N. Hakansson, *Portfolio analysis.* W.W. Norton, New York, 1997.

I. Karatzas, Adaptive control of a diffusion to a goal and a parabolic Monge-Ampère type equation, *Asian J. Mathematics* **1** (1997), 324-341.

I. Karatzas, S.E. Shreve, *Brownian motion and stochastic calculus*, 2nd ed., Springer-Verlag, New York, 1991.

I. Karatzas, S.E. Shreve, *Methods of mathematical finance*, Springer-Verlag, New York, 1998.

I. Karatzas, X.-X. Xue, A note on utility maximization under partial observations, *Mathematical Finance*, **1** (1991), no. 2, 57-70.

I. Karatzas, X. Zhao, Bayesian adaptive portfolio optimization, *Handbook of Mathematical Finance*, Cambridge University Press, 632-670, 2001.

Y. Kuwana, Certainty equivalence and logarithmic utilities in consumption/investment problems, *Mathematical Finance* **5** (1995), 297-310.

P. Lakner, Utility maximization with partial information. *Stoch. Processes Appl.* **56** (1995), 247-273.

P. Lakner, Optimal trading strategy for an investor: the case of partial information. *Stoch. Processes Appl.*, **76** (1998), 77-97.

R.S. Liptser, A.N. Shiryaev, *Statistics of random processes. I. General theory*, Berlin, Heidelberg, New York: Springer-Verlag (2nd ed), 2000.

A.W. Lo, Maximum likelihood estimation of generalized Itô processes with discretely sampled data, *Econometrics Theory* **4** (1988), 231-247.

R. Merton, Lifetime portfolio selection under uncertainty: the continuous-time case. *Rev. Econom. Statist.* **51** (1969), 247-257.

N.D. Pearson, T.-S. Sun, Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingresoll, and Ross model, *J. Finance* **49** (1994), 1279-1304.

R.W. Rishel, Optimal portfolio management with partial observations and power utility function. In *Stochastic Analysis, Control, Optimization and Applications.* (W.M. McEneaney, G. Yin, Q. Zhang, eds.), Birkhauser, 1999, 605-620.

B.L. Rozovskii, *Stochastic evolution systems. Linear theory and applications to non-linear filtering.* Kluwer Academic Publishers. Dordrecht-Boston-London, 1990.

J.T. Williams, Capital asset prices with heterogeneous beliefs, *J. Financial Economics* **5** (1977), 219-240.

J. Yong and X. Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations.* Springer-Verlag, New York. 1999.